

# Deep Learning in Historical Document Image Analysis and Recognition

การเรียนรู้เชิงลึกในการวิเคราะห์และรับจำ  
รูปภาพเอกสารโบราณ

สรายุทธ กรวัตร์ตัน

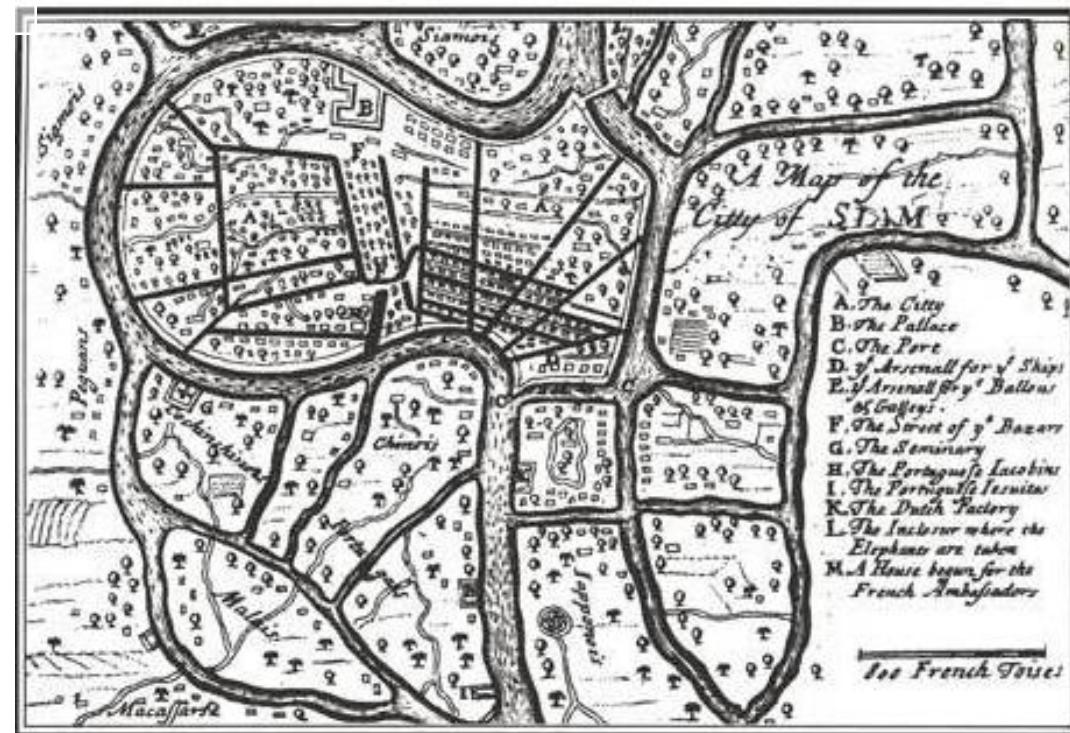
# Historical Documents

- Contain important historical information about a person, place, or event
- Historical Documents etc.
  - Bark manuscript
  - Palm Leaf
  - Old books
- Digital Libraries and Archives
  - Automatic System



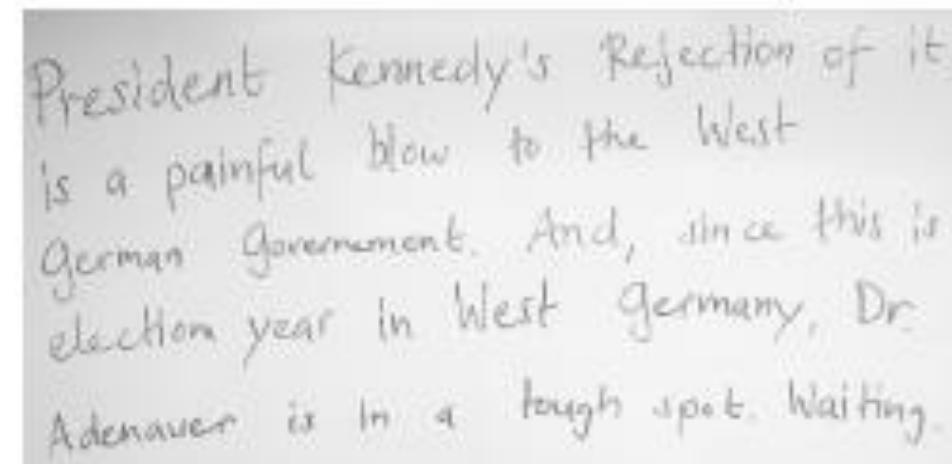
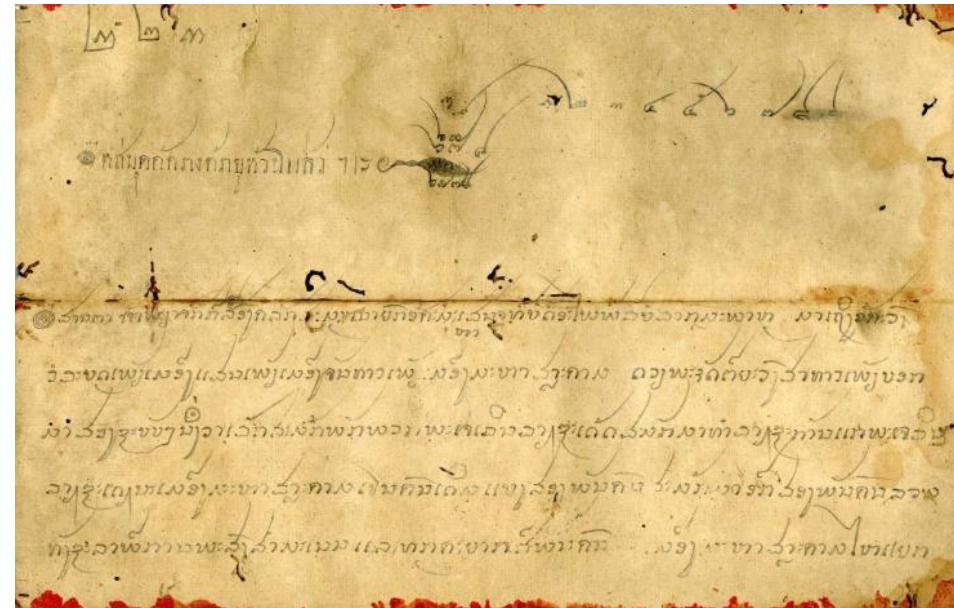
# Document Image Analysis and Recognition

- Document Image Analysis and Recognition DIAR
  - Forms ,bank check and financial document
- Understanding of the document contexts
- Automatic Reading System
- Automatic Information Extraction
  - Information Retrieval
  - Layout text or image extraction, graphical drawing



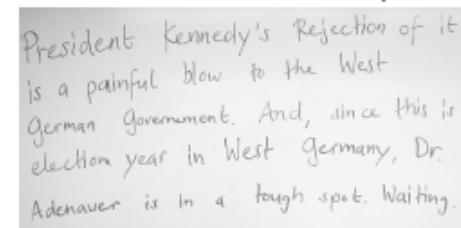
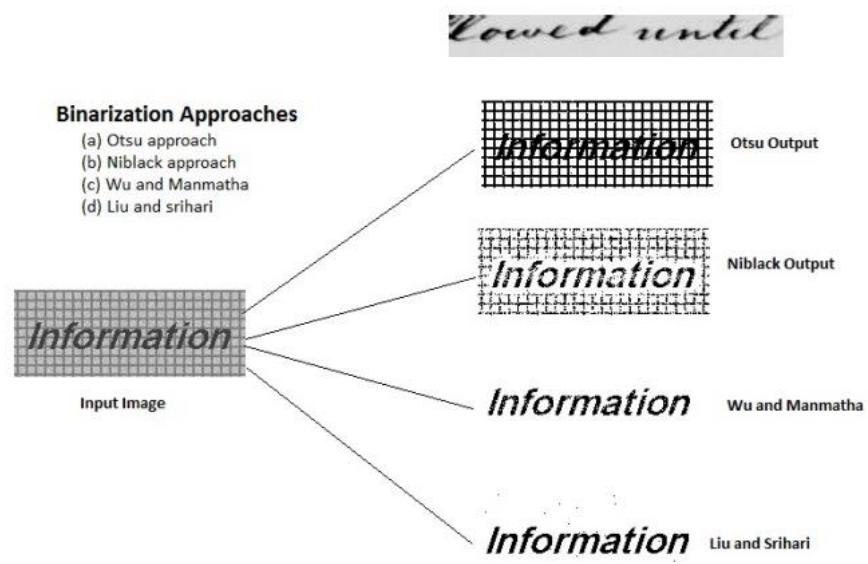
# Handwritten Document

- Historical vs Modern Document
  - Quality
  - Handwritten style
- Printed-Text
  - Optical Character Recognition OCR
  - achieves high-quality Result
- Handwriting ,
  - variability in writing style or cursive text
  - Online Handwriting Recognition ,
    - Pen pressure , stroke and list of points
  - Offline Handwriting Recognition

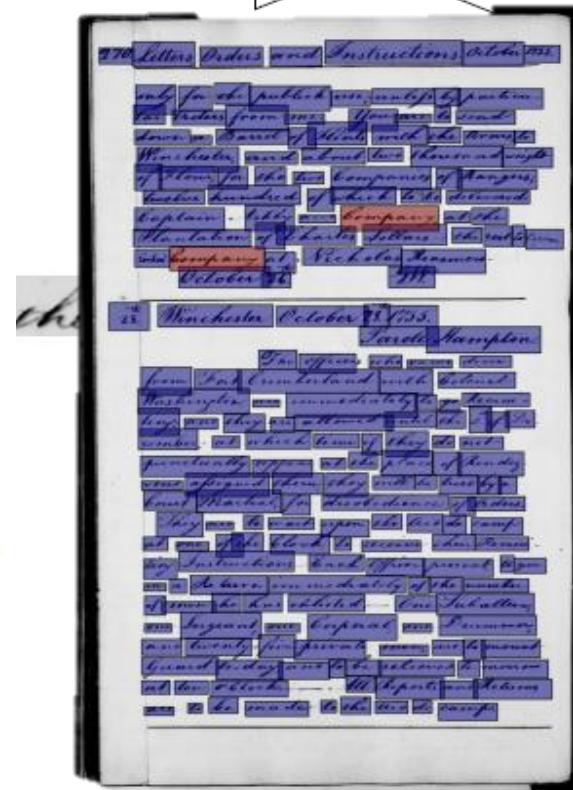


# Research Problem in Historical Document

- Handwritten Character Recognition
  - Classification Document
  - Document Processing
  - Word Spotting
  - Word Recognition



*President Kennedy is Rejection of it  
is a painful blow to the West  
German Government And since this is  
selection year in West Germany of  
Adenauer is in a tough spot waiting*



# Handwritten Character Recognition

- Feature Extraction or Descriptor
- Machine Learning Algorithm

O. Surinta, M.F. Karaaba, T.K. Mishra, L.R.B. Schomaker and M.A. Wiering (2015).

Recognizing Handwritten Characters with Local Descriptors and Bags of Visual Words.

O. Surinta, M.F. Karaaba, L.R.B. Schomaker and M.A. Wiering (2015). Recognition of handwritten characters using local gradient feature descriptors

P. Inkeaw, P. Charoenkwan, H. L. Huang, S. Marukata, S, Y Ho and J. Chaijaruwanich (2017). Recognition of Handwritten Lanna Dhaamma Characters using a set of optimally designed moment feautres



# Classification Documents

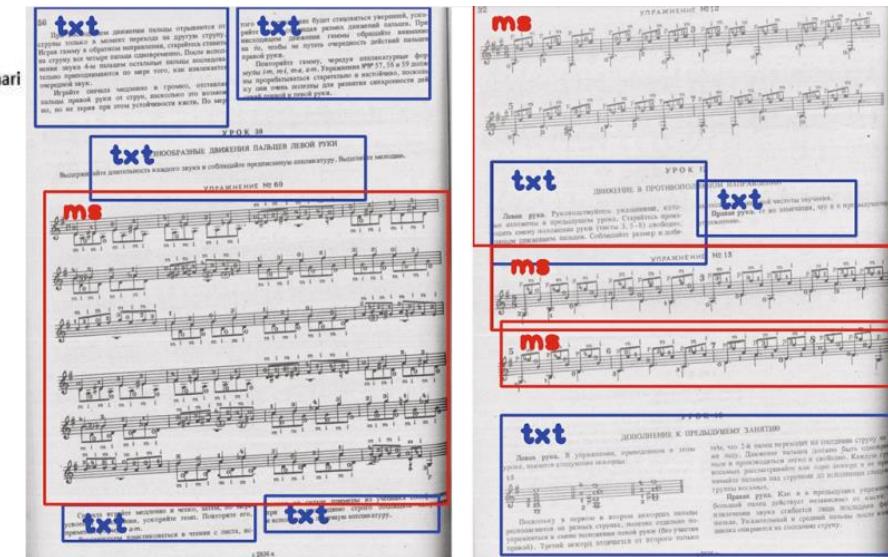
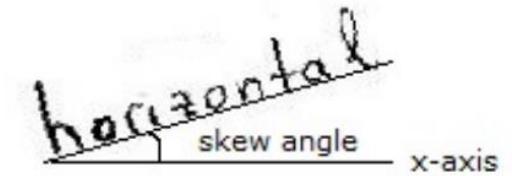
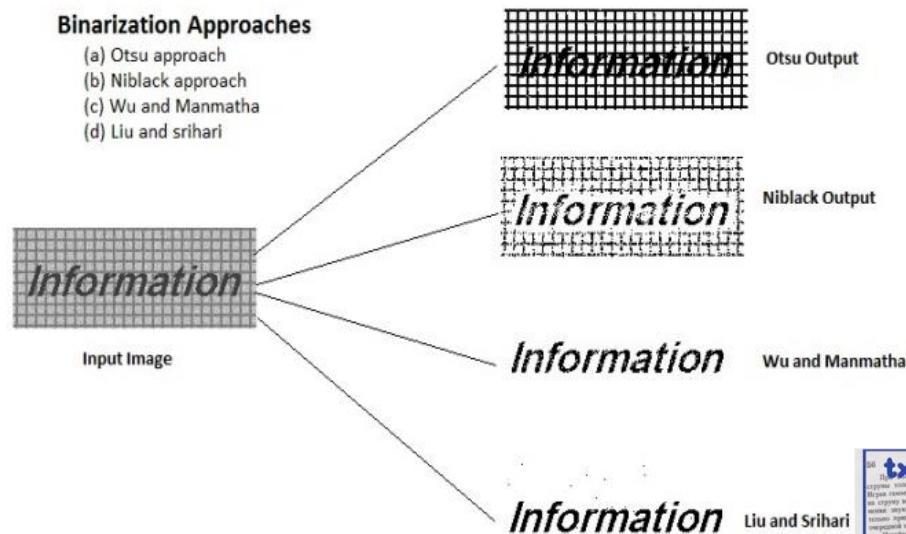
- Dating and Localization
- Writer Identification



He et al. (2016) He S., Samara P., Burgers J., Schomaker L. **Writer identification using curvature-free features** Pattern Recognition  
S. He, (2016) P. Samara, J. Burgers, L. Schomaker **A multiple-label guided clustering algorithm for historical document dating and localization** IEEE Trans. Image Processing

# Document Processing (1)

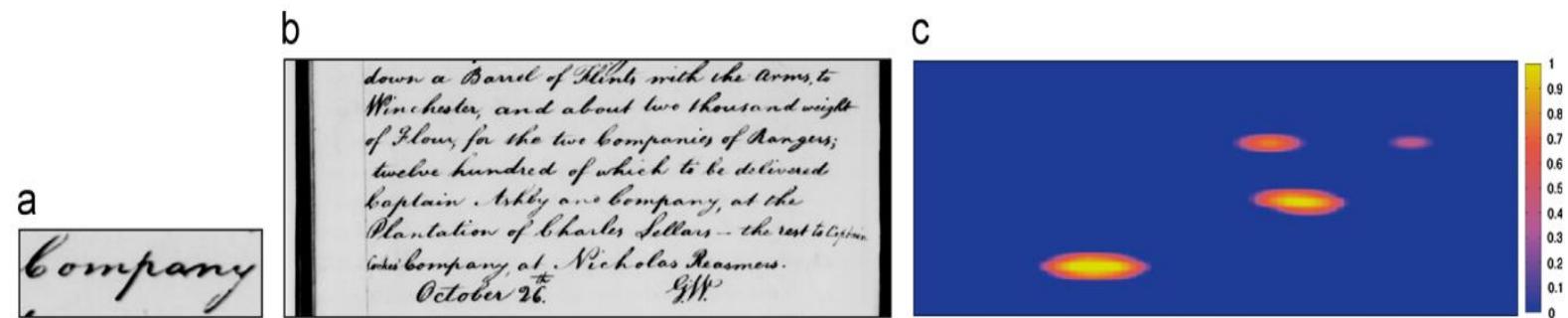
- Binarization
- Normalization
- Segmentation



Subhash P., Neeta N.: Handwritten Text Documents Binarization and Skew Normalization Approaches , International Conference on Intelligent Human Computer Interaction

# Document Processing (1)

- Segmentation – free
- Binarization –free



Marçal Rusiñol , David Aldavert, Ricardo Toledo, Josep Lladós Efficient segmentation-free keyword spotting in historical document collections Int. Pattern Recognition, 48(2015), pp.545–555

M. R. Yousefi, M. R. Soheili, T. M. Breuel†, E. Kabir and D. Stricker  
Binarization-free OCR for Historical Documents Using LSTM  
Networks

GT:	Mann	eintrag.
pred:	Mann	eintrag.

# Word Spotting

- If transcription is not accomplish
- Information Retrieval System
  - Special case of content -based image retrieval
- Query type
  - Query by Example
  - Query by String

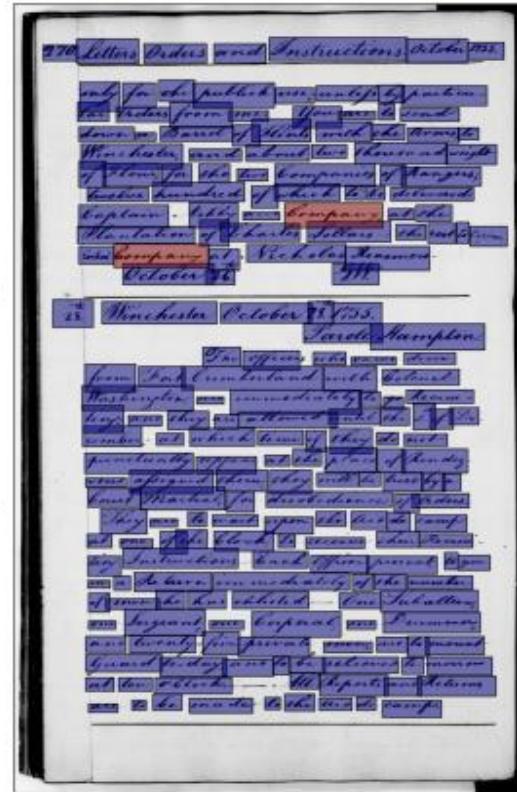
Query by Example

A handwritten word 'Company' is shown in a rectangular box.

Query by String

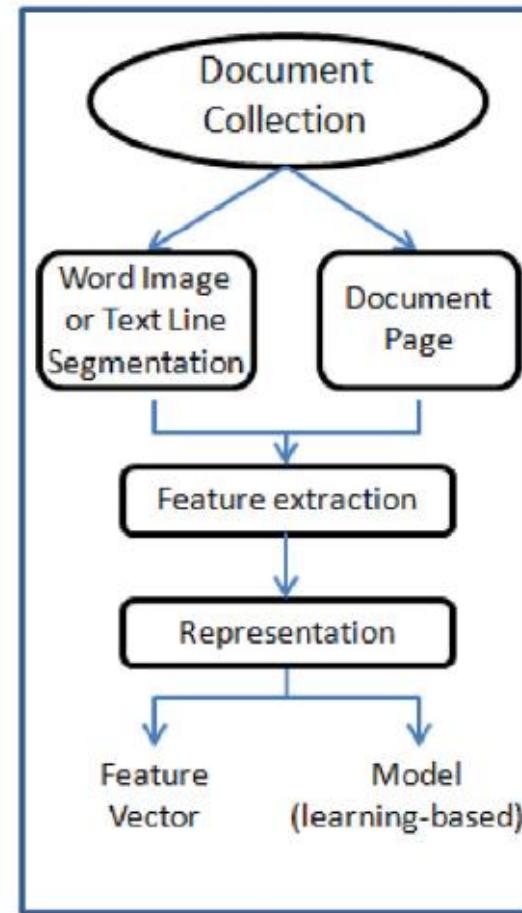
A printed word 'Company' is shown in a rectangular box.

Image by Fink, ICFHR 2018



# Word Spotting System Architecture

- Preprocessing
  - Binarization
  - Segmentation , line word
  - Normalization , skew and slat
- Feature Extraction
  - Histograms of Gradients HoG
  - Scale Invariant Feature Transform SIFT , SUFT
- Representation
  - Fixed-length vector
  - Bag of Visual Word (BoVW)
- Matching Process
  - Dynamics Time Warping DTW
  - Euclidean Distance
  - Cosine Distance



# Word Recognition

- Sequence Technique ,
  - Hidden Markov Models
  - LSTM RNN

A. El-Yacoubi, M. Gilloux, R. Sabourin, and C.Y. Suen, Fello. (199):

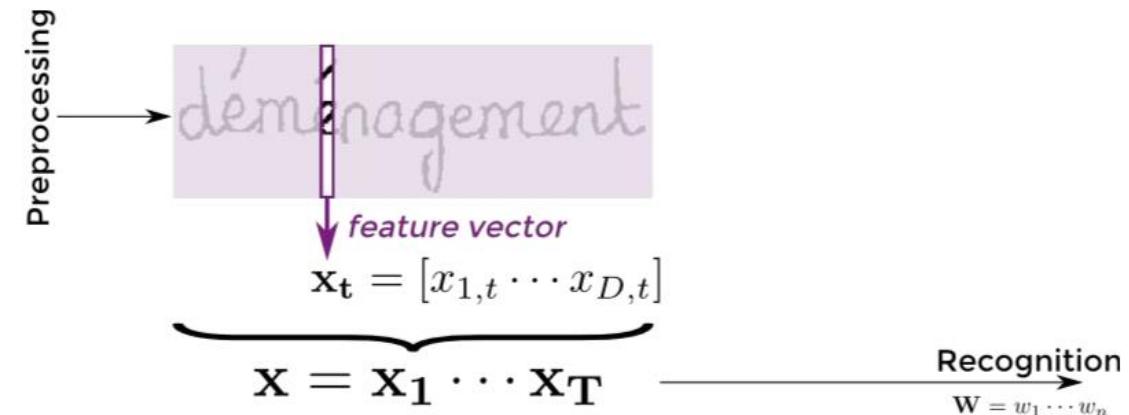
An HMM-Based Approach for Off-Line Unconstrained Handwritten Word  
Modeling and Recognition.

IEEE Transactions on Pattern Analysis and Machine Intelligence

Almaz' an, J., Gordo, A., Forn' es, A., Valveny, E. (2014):

Word Spotting and Recognition with Embedded Attributes.

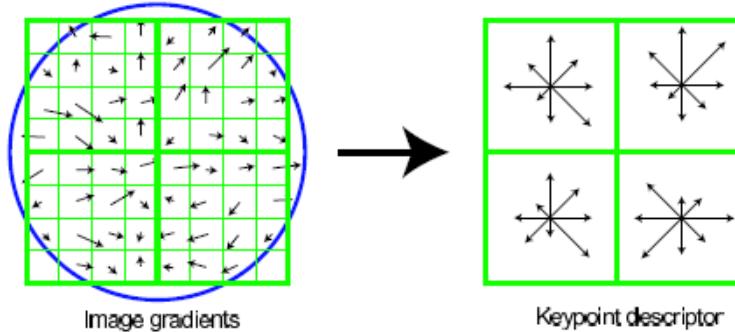
IEEE Transactions on Pattern Analysis and  
Machine Intelligence



# Background Knowledge

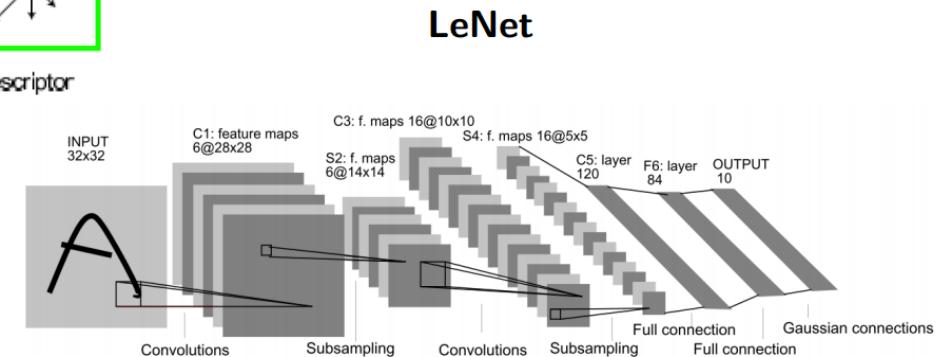
- Feature Extraction

- SIFT
- HOG
- LBP



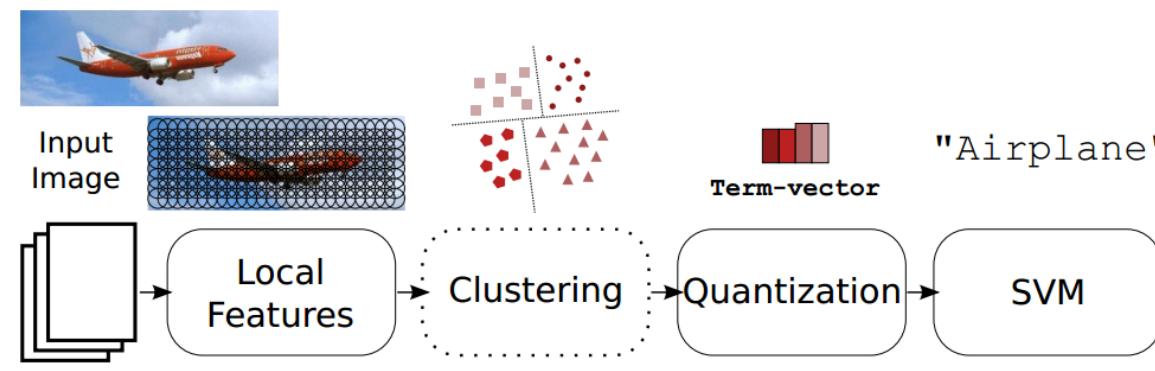
- Machine Learning

- SVM
- MLP
- BoVW



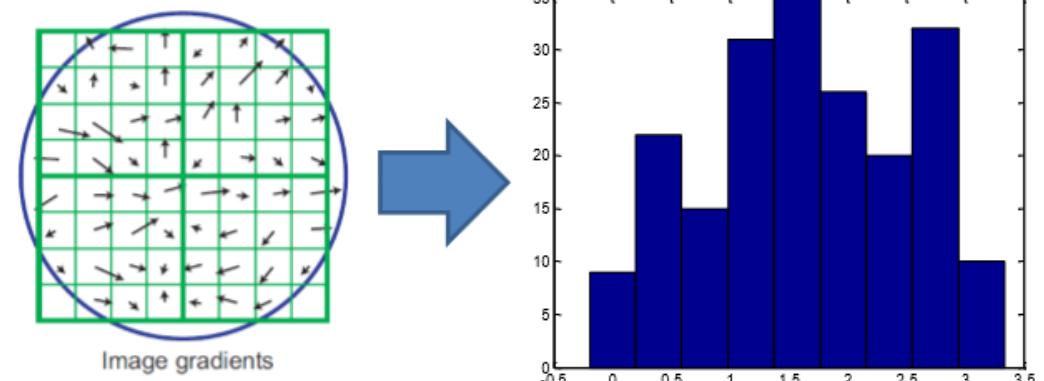
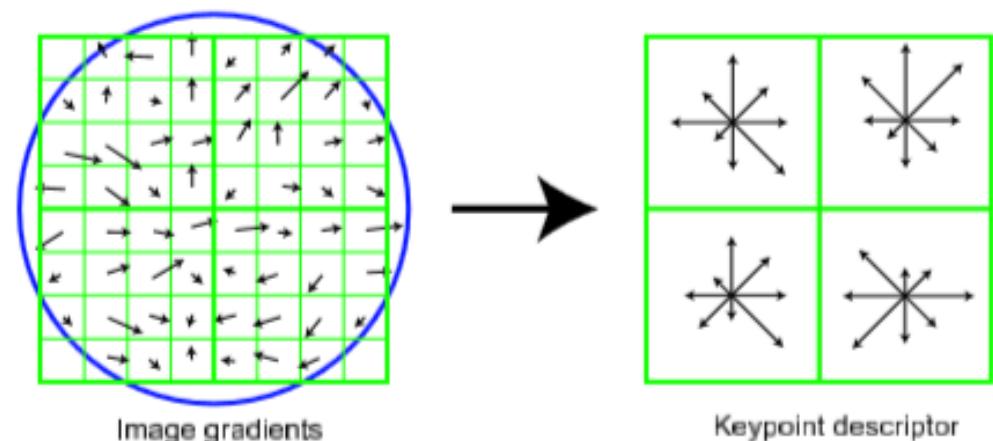
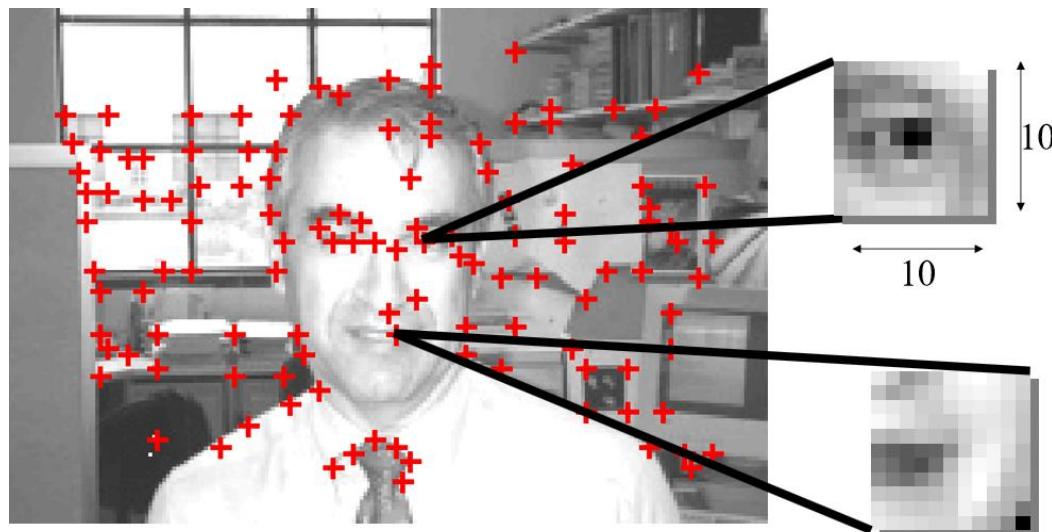
- Deep Learning

- CNN
- FCN
- LSTM



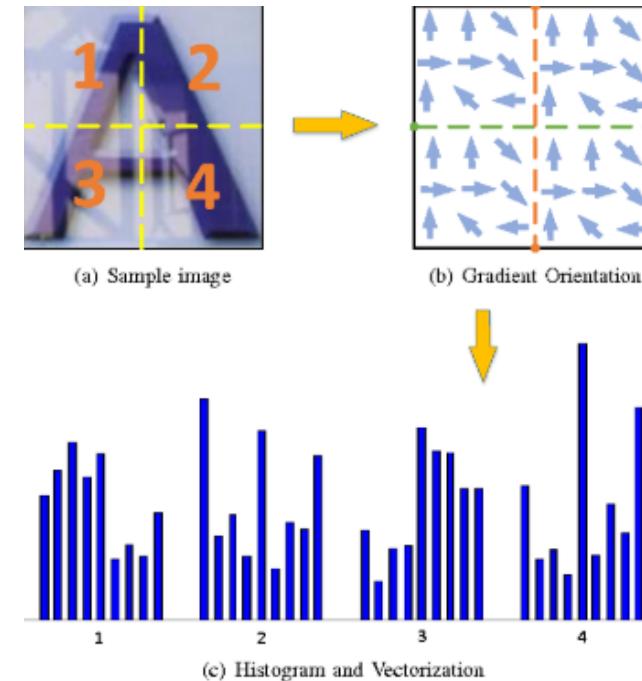
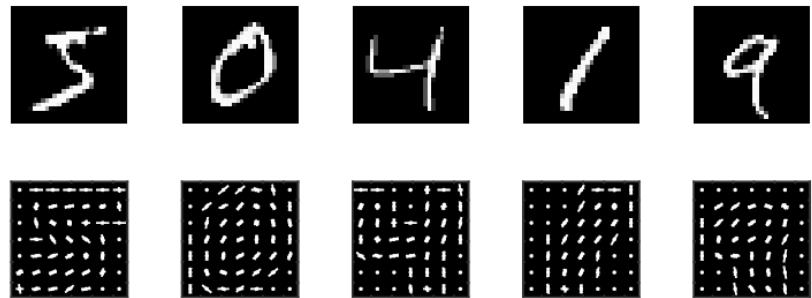
# Feature Extraction (1)

- Scale Invariant Feature Transform: SIFT

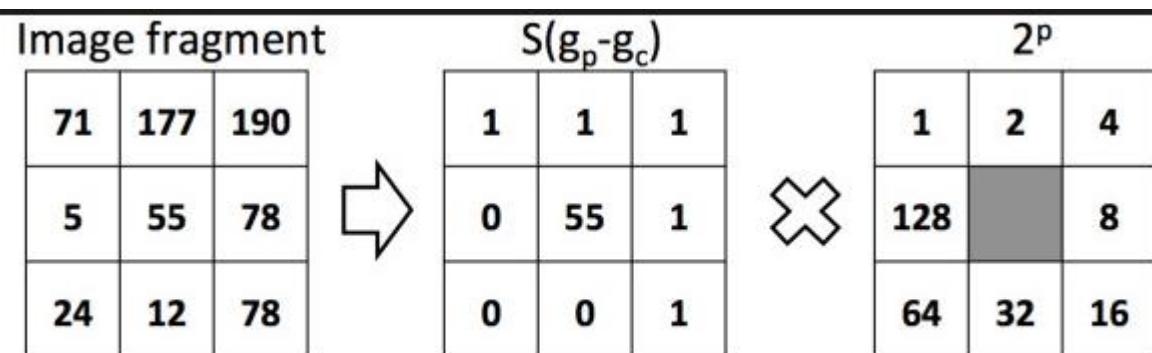


# Feature Extraction (2)

- Histogram of Oriented Gradients: HOG

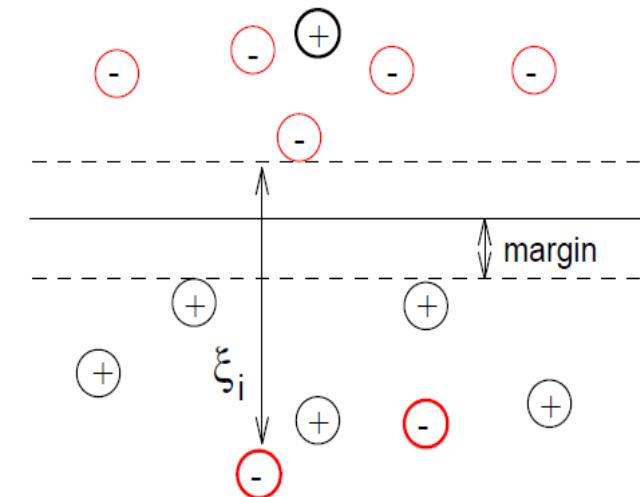
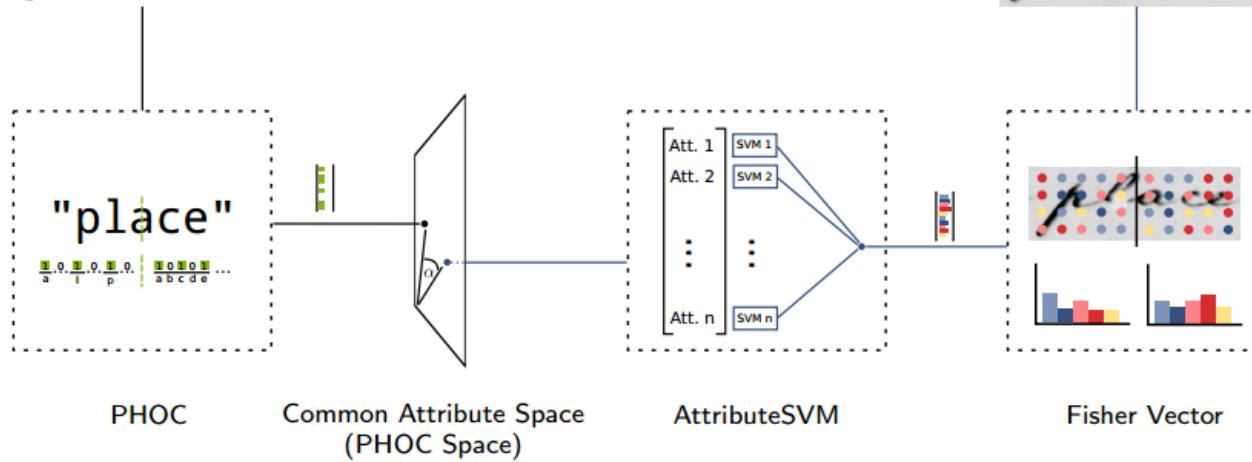


- Local Binary Pattern: LBP

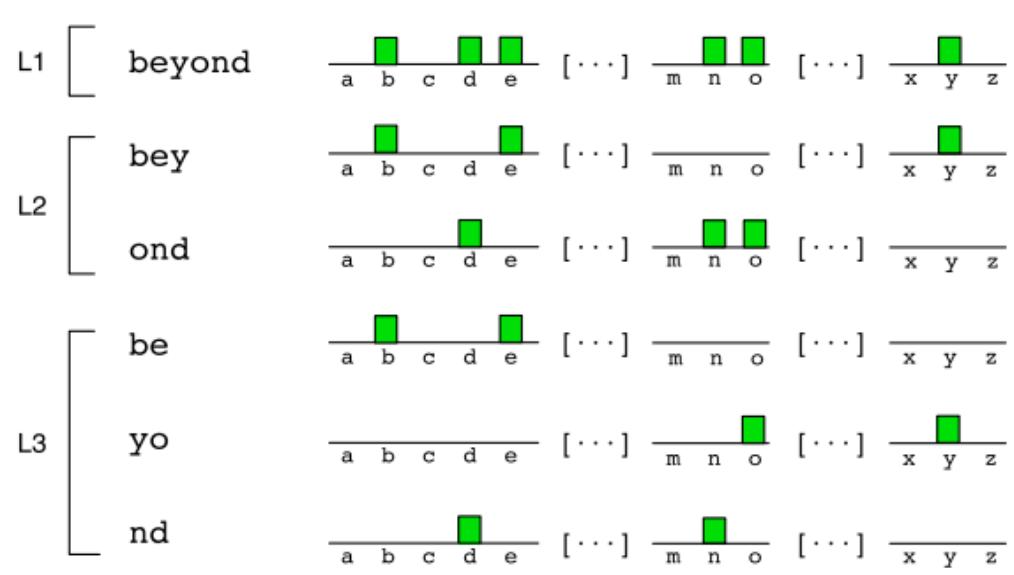


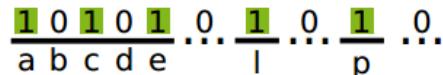
# Support Vector Machine

"place"

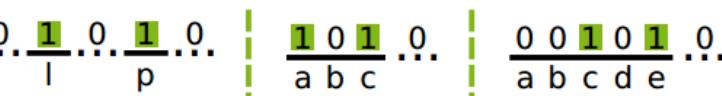


# Pyramidal Histogram of Characters



Level 1 "place" 

Level 2 "place" 

Level 3 "place" 

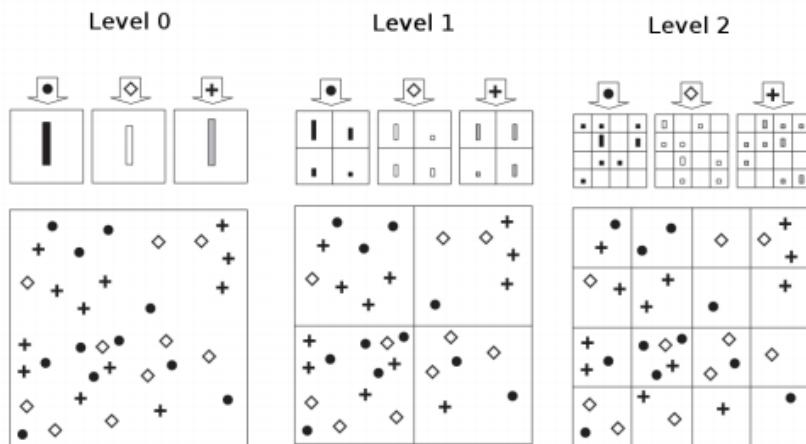
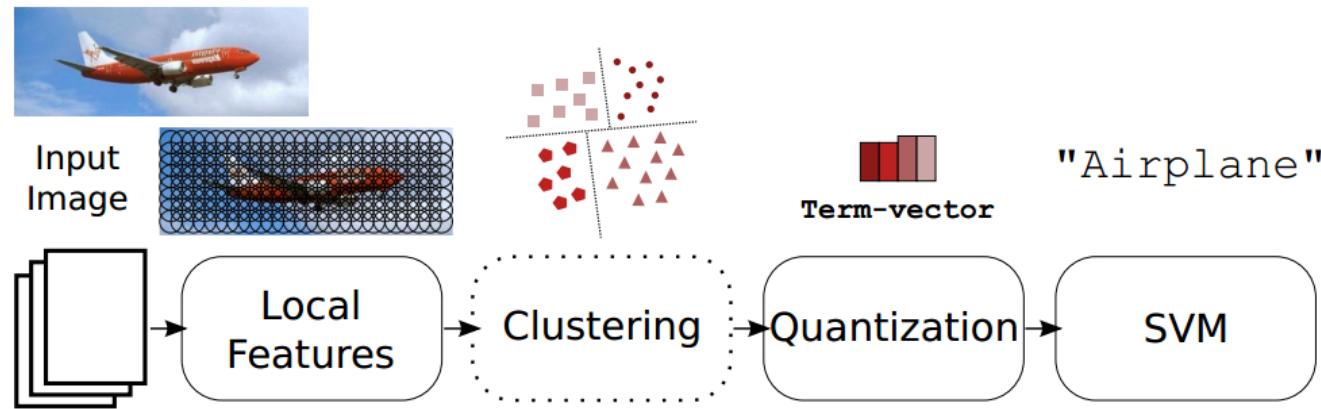
$$604 \text{ Attributes} = 100 + 504 = 604$$

$$50 \text{ well-known bigram level 2 } 2 \times 50 = 100$$

$$36 \times (2+3+4+5) = 504$$

# Bag of Visual Word

## BoF: Bag of Features

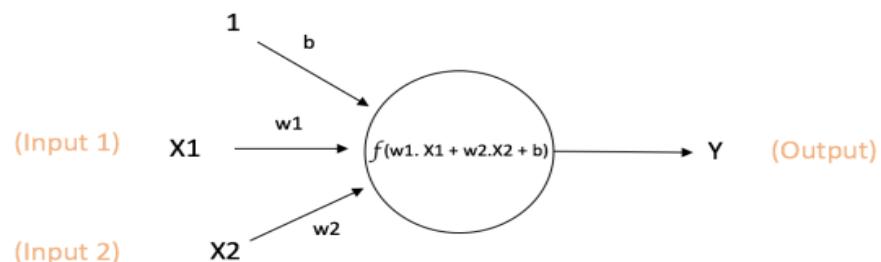
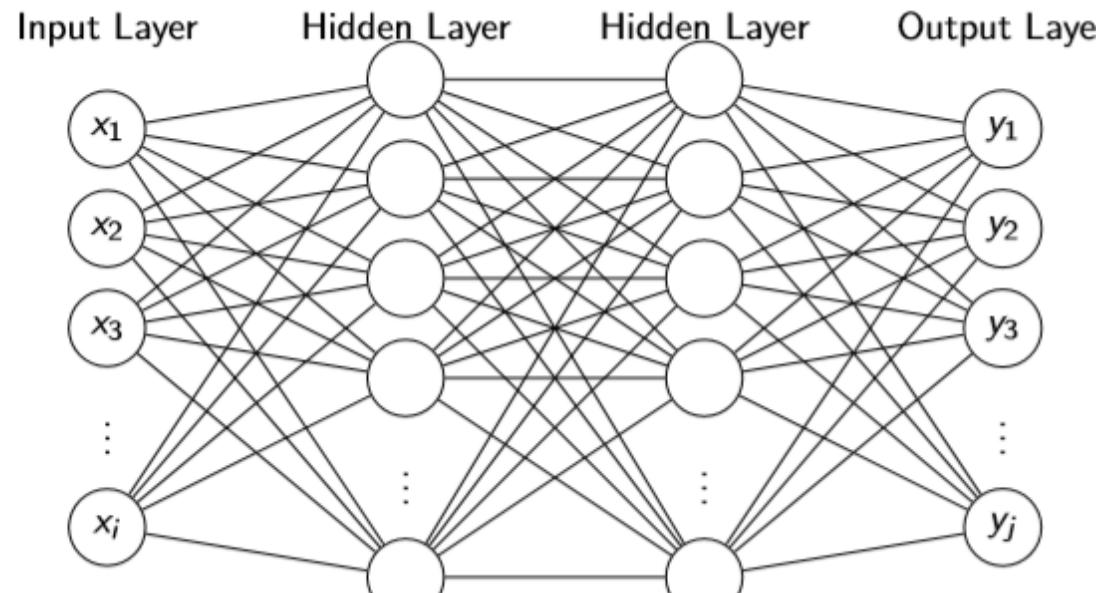


## Spatial-pyramid Model

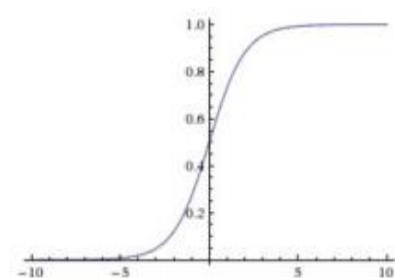
Lazebnik, S., Schmid, C., Ponce, J.: **Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories**, IEEE Conference on Computer Vision and Pattern Recognition, 2006.

# Artificial Neural Network

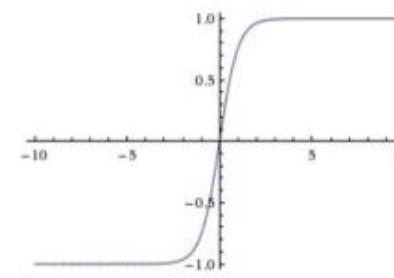
- Multilayer Perceptron



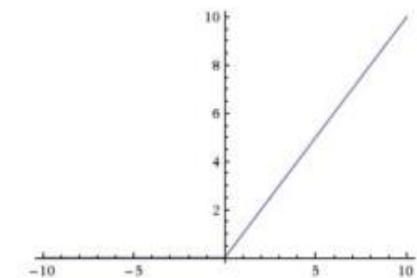
$$\text{Output of neuron} = Y = f(w_1 \cdot X_1 + w_2 \cdot X_2 + b)$$



Sigmoid



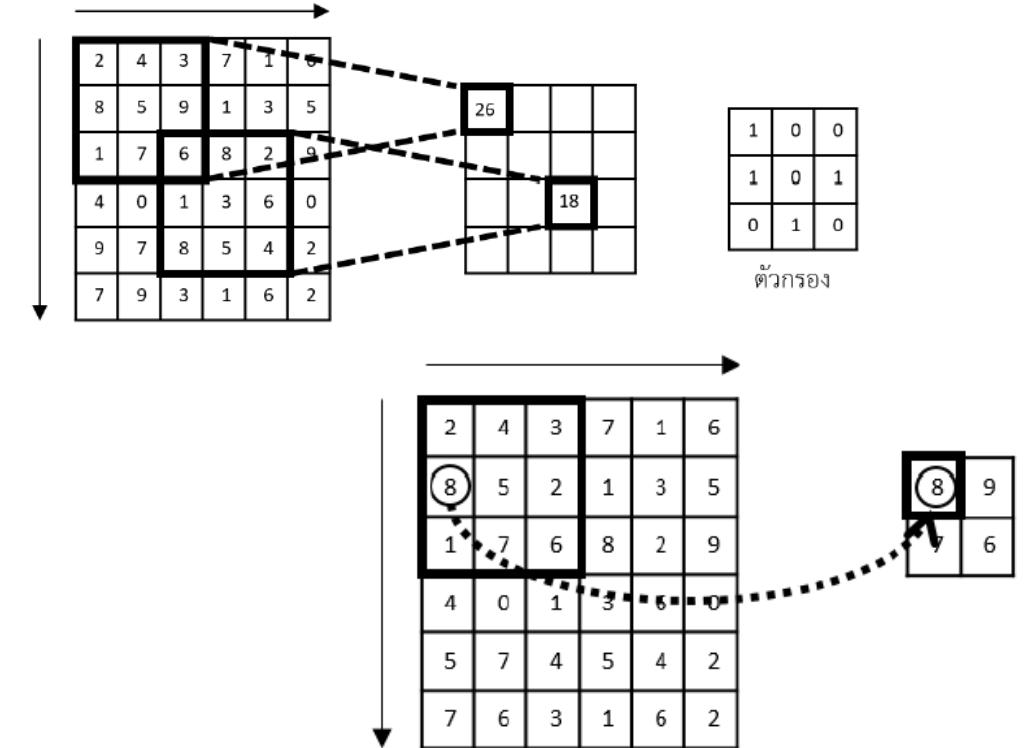
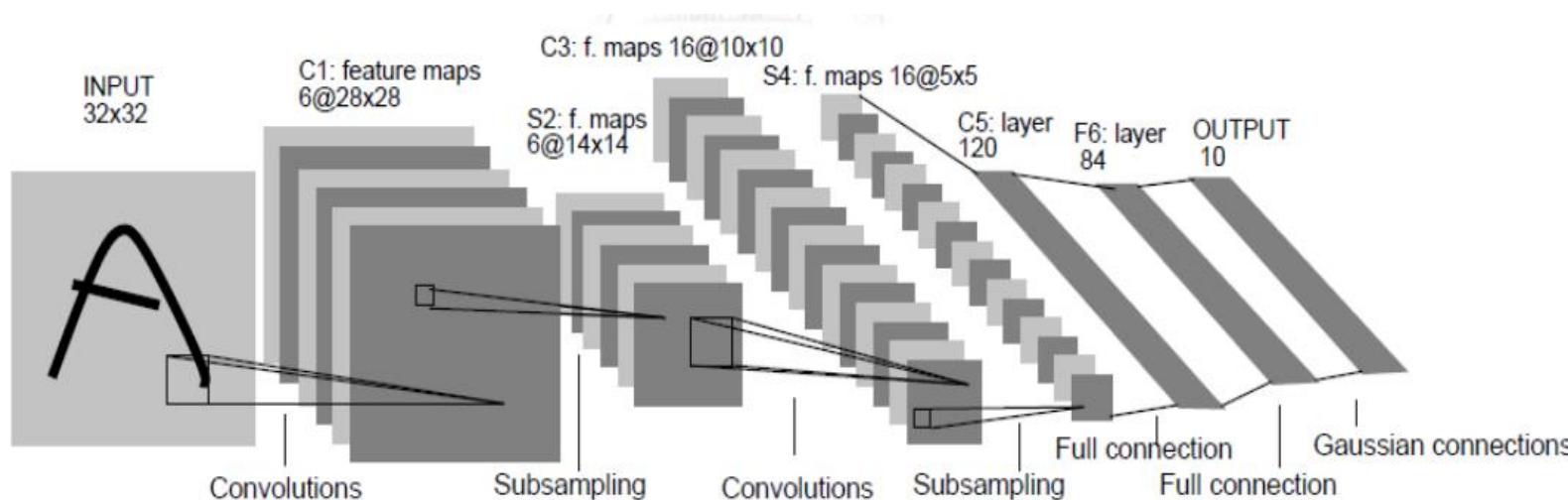
tanh



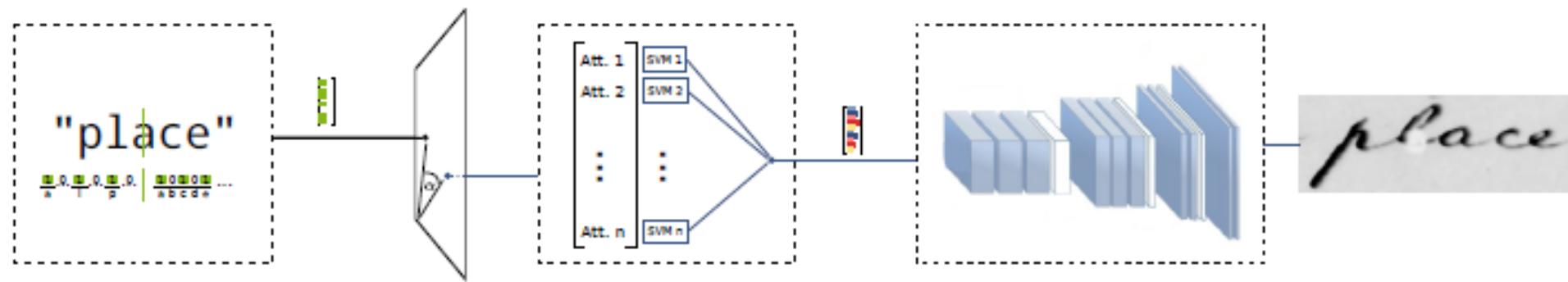
ReLU

# Deep Learning

- Convolutional Neural Network CNN
  - Convolution Layer
  - Max Pooling
  - Fully Connected Layer

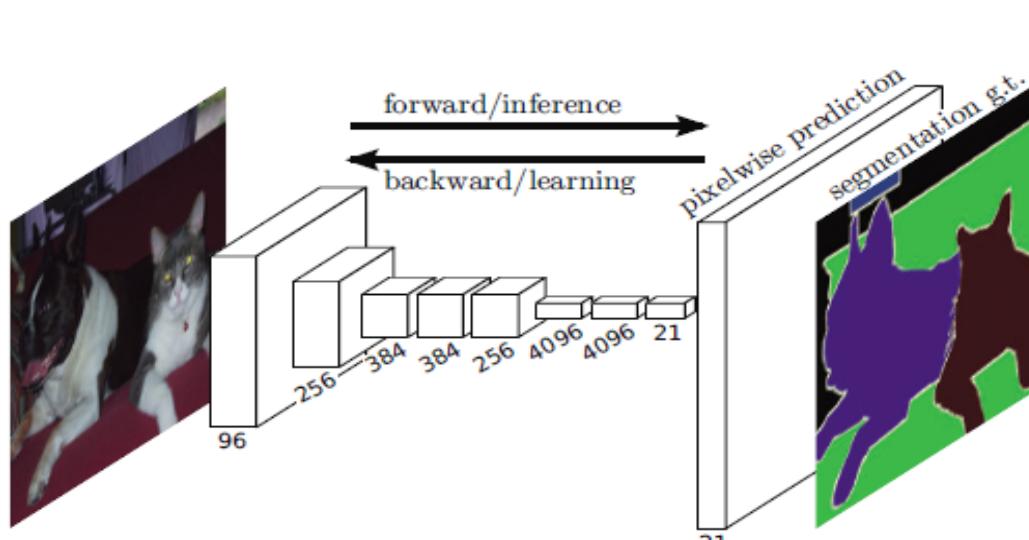
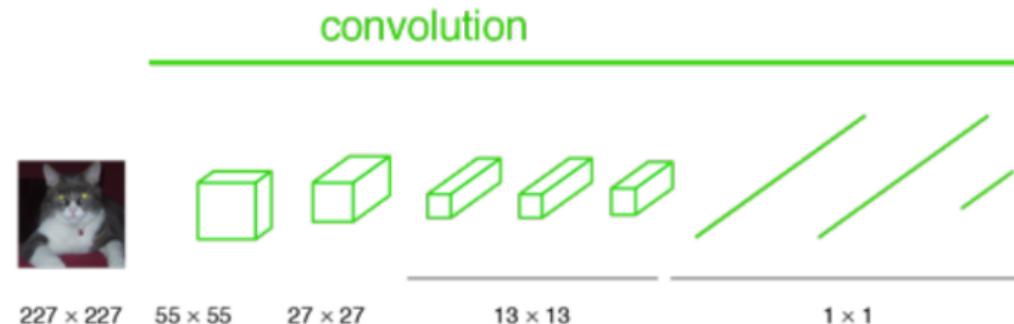


# Convolutional Neural Network



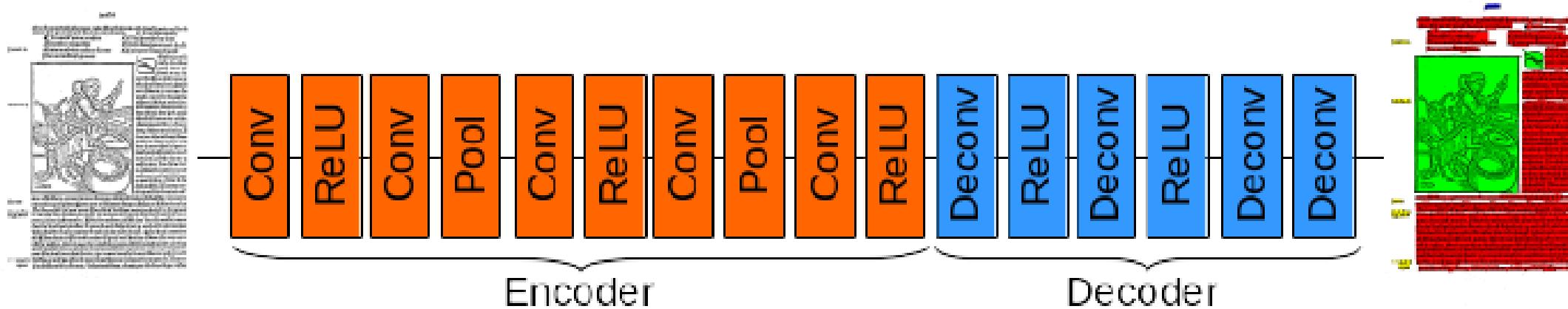
Sebastian S and Gernot A. Fink (2018): Attribute CNNs for Word Spotting in Handwritten Documents. International Journal on Document Analysis and Recognition, 2018

# Fully Convolutional Network



# Deep Learning in Historical Document (2)

- Segmentation FDN

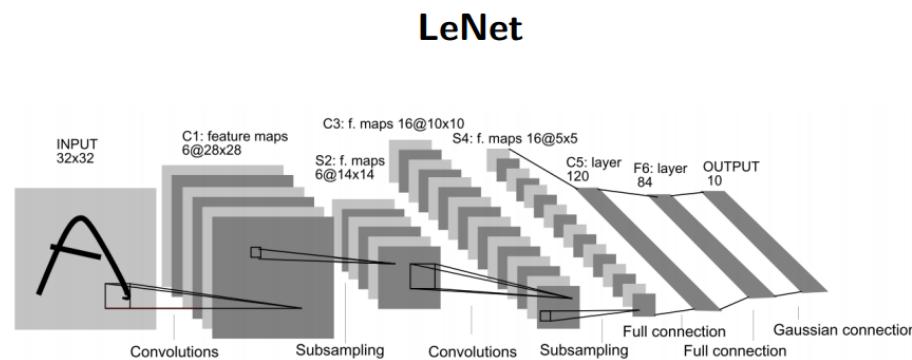


Christoph Wick and Frank Puppe (2018)

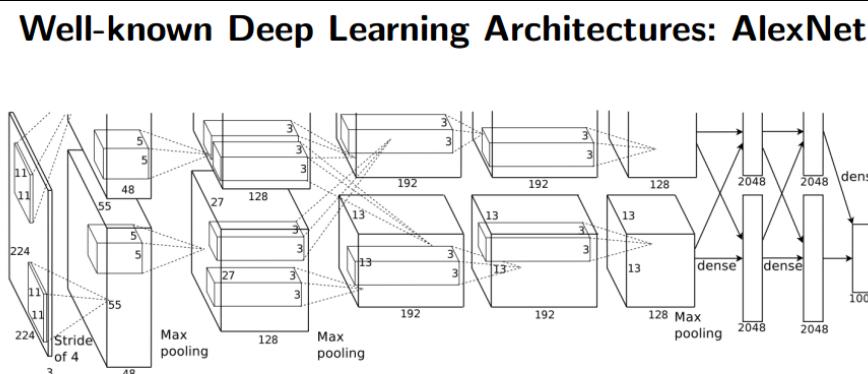
Fully Convolutional Neural Networks for Page Segmentation of Historical Document Images

# Deep learning Model

ResNet, ResNet 101, GooLeNet ,AlexNet, YOLOv2 etc..



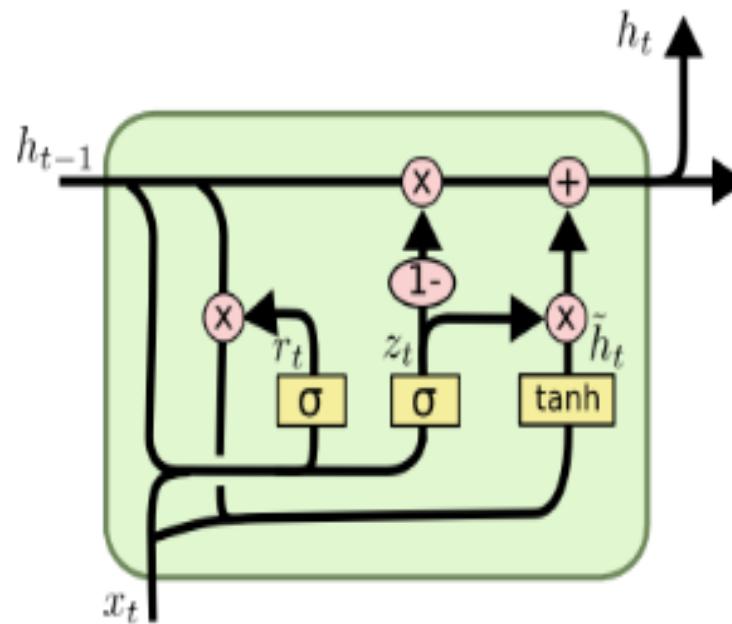
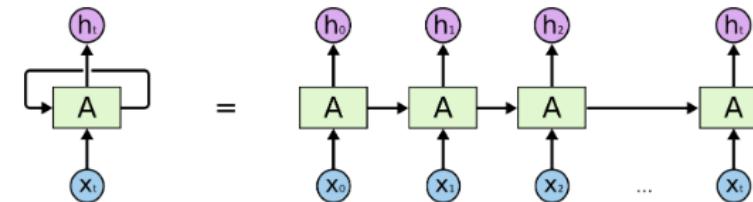
(Source: [Ercihui et al., 1990])



(Source: Krizhevsky et al., 2012)

# Long Short Term Memory LSTM

- Recurrent Neural Network RNN
- Long Short Term Memory LSTM



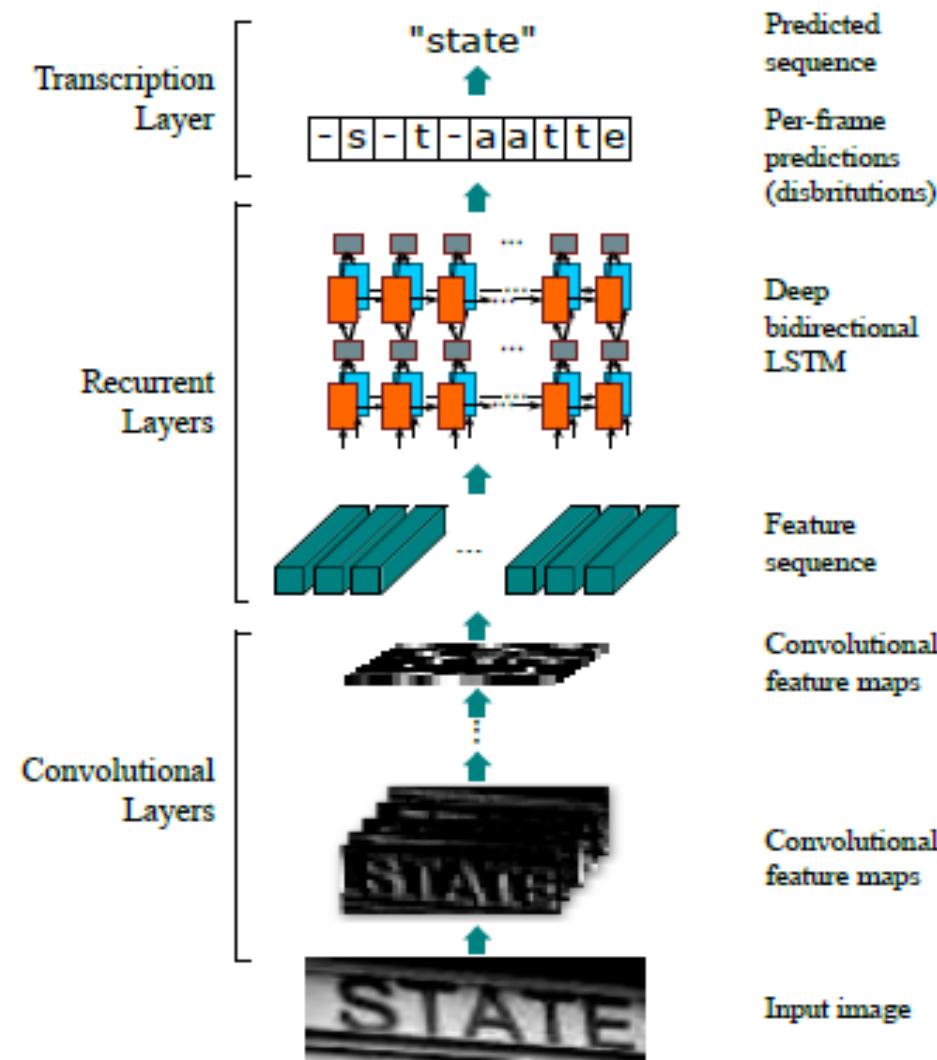
$$z_t = \sigma (W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma (W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh (W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

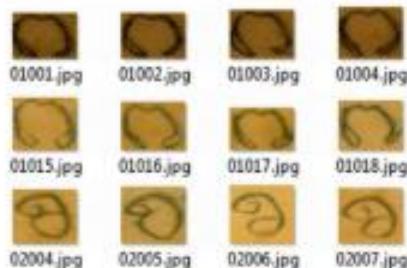
# Long Short Term Memory LSTM



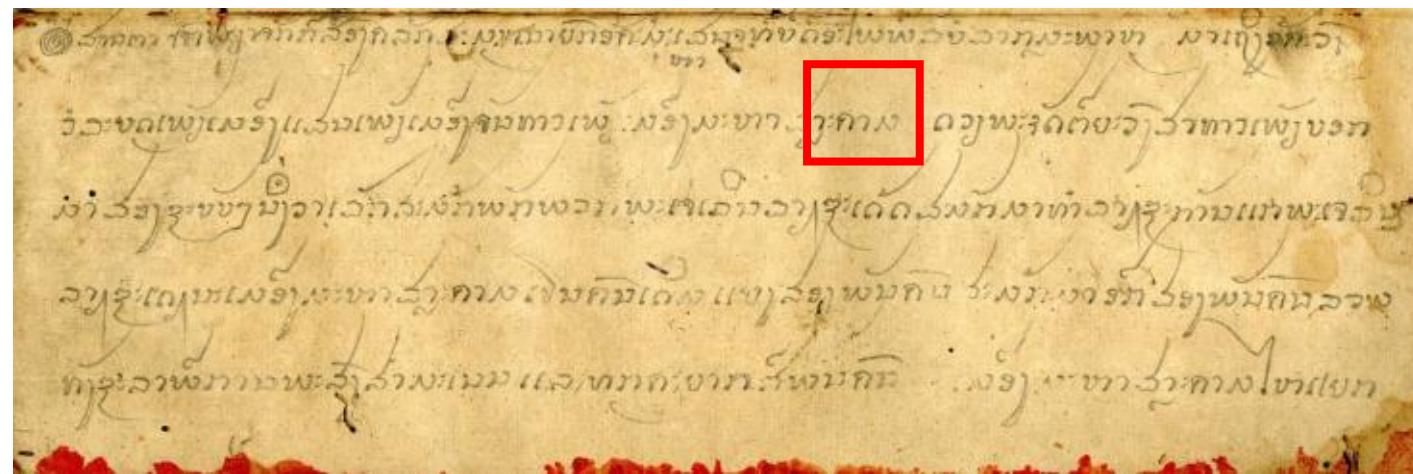
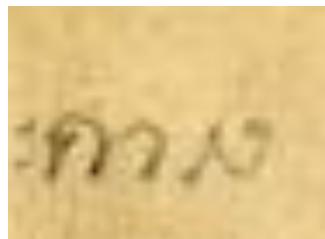
# Future Work (1)

Base on Applied Deep Learning in Historical Document

- Handwritten Character Recognition

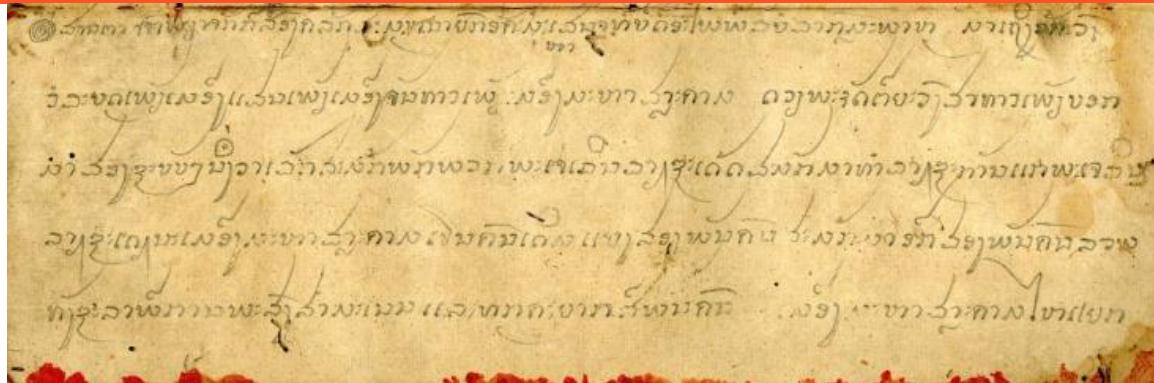


- Word Spotting



# Future Work (2)

- Word Transcription



- ๐๗) สถานดำเนินการจัดกิจกรรมคลังความมุหะมิค อัคคามหาเสนานีบดี อะไฟฟลีบล่า กิลมะ พาหุ มาถึงอัคคะวิง-
  - วัฒนบุคเพี้ยเมืองแสน เพี้ยเมืองจัน ท้าวเพี้ยเมืองมหาสารากาน ด้วยพระขัตติยะวงศ์สากาท้าวเพี้ยบอกร
  - มาสองจะบันๆ นึงว่า เล็กสมัคพักพวง พระเจลินลาสจะเด็ค สมัคกามาทำลาสจะกันแกพระเจลิน-
  - ลาสจะเด็ค นะ เมืองมหาสารากานเป็นคนเดิมแบงสองพันคน สมัคกามาอึกสองพันคน ล้วน
  - ทั้งจะลาพิกาน พะสิ่ง สามะเนน และทุกจะยกสีพันคน แม่อึงมหาสารากานให้แยก
  - ๐๘) สารตราเข้าพระยาจักรีศรีองครักษ์ อัคคามหาเสนานบดีอภัยพิรบรมกรมพาหุ มาถึงอัครวงศ์ วรบุตร เพี้ยเมืองแสน เพี้ยเมืองจัน ท้าวเพี้ยเมืองมหาสารากาน ด้วยพระขัตติยะวงศ์ ท้าวเพี้ยบอกร
  - มาสองจะบัน จบันหนึ่งว่า เล็กสมัคพรรคพวง พระเจริญราชเดช สมัคกามาทำราชการแก่พระเจริญ-
  - ราชเดช ณ เมืองมหาสารากานเป็นคนเดิมแบงเป็นสองพันคน สมัคกามาอึกสองพันคนรวม
  - ทั้งจะรา พิกาน พะสิ่ง สามะเนน และทุกจะยกสีพันคน เมืองมหาสารากานให้แยก